Artificial Intelligence 231 (2016) 1–16

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Voting rules as error-correcting codes ☆

Ariel D. Procaccia, Nisarg Shah *, Yair Zick

Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:
Received 21 January 2015
Received in revised form 13 October 2015
Accepted 20 October 2015
Available online 27 October 2015

Keywords:
Social choice
Voting
Ground truth
Adversarial noise
Error-correcting codes

ABSTRACT

We present the first model of optimal voting under adversarial noise. From this viewpoint, voting rules are seen as error-correcting codes: their goal is to correct errors in the input rankings and recover a ranking that is close to the ground truth. We derive worst-case bounds on the relation between the average accuracy of the input votes, and the accuracy of the output ranking. Empirical results from real data show that our approach produces significantly more accurate rankings than alternative approaches.

1. Introduction

Social choice theory develops and analyzes methods for aggregating the opinions of individuals into a collective decision. The prevalent approach is motivated by situations in which opinions are *subjective*, such as political elections, and focuses on the design of voting rules that satisfy normative properties [1].

An alternative approach, which was proposed by the marquis de Condorcet in the 18th Century, had confounded scholars for centuries (due to Condorcet's ambiguous writing) until it was finally elucidated by Young [34]. The underlying assumption is that the alternatives can be *objectively* compared according to their true quality. In particular, it is typically assumed that there is a ground truth *ranking* of the alternatives. Votes can be seen as noisy estimates of the ground truth, drawn from a specific *noise model*. For example, Condorcet proposed a noise model where — roughly speaking — each voter (hereinafter, *agent*) compares every pair of alternatives, and orders them correctly (according to the ground truth) with probability $p > 1/2$; today an equivalent model is attributed to Mallows [26]. Here, it is natural to employ a voting rule that always returns a ranking that is *most likely* to coincide with the ground truth, that is, the voting rule should be a *maximum likelihood estimator (MLE)*.

Although Condorcet could have hardly foreseen this, his MLE approach is eminently applicable to crowdsourcing and human computation systems, which often employ voting to aggregate noisy estimates; EteRNA [24] is a wonderful example, as explained by Procaccia et al. [30]. Consequently, the study of voting rules as MLEs has been gaining steam in the last decade [16,15,19,33,32,25,30,2–4,27,12,13].

Despite its conceptual appeal, a major shortcoming of the MLE approach is that the MLE voting rule is specific to a noise model, and that noise model — even if it exists for a given setting — may be difficult to pin down [27]. Caragiannis et al. [12,13] have addressed this problem by relaxing the MLE constraint: they only ask that the probability of the voting

☆ A preliminary version of the paper appears in the Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI), 2015.

* Corresponding author.
E-mail addresses: arielpro@cs.cmu.edu (A.D. Procaccia), nkshah@cs.cmu.edu (N. Shah), yairzick@cmu.edu (Y. Zick).

http://dx.doi.org/10.1016/j.artint.2015.10.003

rule returning the ground truth go to one as the number of votes goes to infinity. This allows them to design voting rules that uncover the ground truth in a wide range of noise models; however, they may potentially require an infinite amount of information.

Our approach. In this paper, we propose a fundamentally different approach to aggregating noisy votes. Instead of assuming probabilistic noise, we assume a known upper bound on the "total noise" in the input votes, and allow the input votes to be *adversarial* subject to the upper bound. We emphasize that in potential application domains there is no adversary that actively inserts errors into the votes; we choose an adversarial error model to be able to correct errors even in the *worst case*. This style of worst-case analysis — where the worst case is assumed to be generated by an adversary — is prevalent in many branches of computer science, e.g., in the analysis of online algorithms [10], and in machine learning [23,9].

We wish to design voting rules that do well in this worst-case scenario. From this viewpoint, our approach is closely related to the extensive literature on *error-correcting codes*. One can think of the votes as a *repetition code*: each vote is a transmitted noisy version of a "message" (the ground truth). The task of the "decoder" is to correct adversarial noise and recover the ground truth, given an upper bound on the total error. The question is: how much total error can this "code" allow while still being able to recover the ground truth?

In more detail, let d be a distance metric on the space of rankings. As an example, the well-known *Kendall tau (KT)* distance between two rankings measures the number of pairs of alternatives on which the two rankings disagree. Suppose that we receive n votes over the set of alternatives $\{a, b, c, d\}$, for an even n, and we know that the average KT distance between the votes and the ground truth is at most $1/2$. Can we always recover the ground truth? No: in the worst-case, exactly $n/2$ agents swap the two highest-ranked alternatives and the rest report the ground truth. In this case, we observe two distinct rankings (each $n/2$ times) that only disagree on the order of the top two alternatives. Both rankings have an average distance of $1/2$ from the input votes, making it impossible to determine which of them is the ground truth.

Let us, therefore, cast a larger net. Inspired by *list decoding* of error-correcting codes (see, e.g., [20]), our main research question is:

> *Fix a distance metric d. Suppose that we are given n noisy rankings, and that the average distance between these rankings and the ground truth is at most t. We wish to recover a ranking that is guaranteed to be at distance at most k from the ground truth. How small can k be, as a function of n and t?*

Our results. We observe that for any metric d, one can always recover a ranking that is at distance at most $2t$ from the ground truth, i.e., $k \le 2t$. We also show that one can pick, in polynomial time, a ranking from the given noisy rankings that provides a weaker $3t$ upper bound. We complement the upper bounds by providing a lower bound of (roughly) $k \ge t/2$ that holds for every distance metric. We also show that an extremely mild assumption on the distance metric improves the lower bound to (roughly) $k \ge t$. In addition, we consider the four most popular distance metrics used in the social choice literature, and prove a tight lower bound of (roughly) $k \ge 2t$ for each metric. This lower bound is our main theoretical result; the construction makes unexpected use of Fermat's Polygonal Number Theorem.

The worst-case optimal voting rule in our framework is defined with respect to a *known* upper bound t on the average distance between the given rankings and the ground truth. However, we show that the voting rule which returns the ranking minimizing the total distance from the given rankings — which has strong theoretical support in the literature — serves as an approximation to our worst-case optimal rule, irrespective of the value of t. We leverage this observation to provide theoretical performance guarantees for our rule in cases where the error bound t given to the rule is an underestimate or overestimate of the tightest upper bound.

Finally, we test our worst-case optimal voting rules against many well-known voting rules, on two real-world datasets [27], and show that the worst-case optimal rules exhibit superior performance as long as the given error bound t is a reasonable overestimate of the tightest upper bound.

Related work. Our work is related to the extensive literature on error-correcting codes that use permutations (see, e.g., [5], and the references therein), but differs in one crucial aspect. In designing error-correcting codes, the focus is on two choices: i) the *codewords*, a subset of rankings which represent the "possible ground truths", and ii) the *code*, which converts every codeword into the message to be sent. These choices are optimized to achieve the best tradeoff between the number of errors corrected and the rate of the code (efficiency), while allowing unique identification of the ground truth. In contrast, our setting has fixed choices: i) every ranking is a possible ground truth, and ii) in coding theory terms, our setting constrains us to the repetition code. Both restrictions (inevitable in our setting) lead to significant inefficiencies, as well as the impossibility of unique identification of the ground truth (as illustrated in the introduction). Our research question is reminiscent of coding theory settings where a bound on adversarial noise is given, and a code is chosen with the bound on the noise as an input to maximize efficiency (see, e.g., [21]).

List decoding (see, e.g., [20]) relaxes classic error correction by guaranteeing that the number of possible messages does not exceed a small quota; then, the decoder simply lists all possible messages. The motivation is that one can simply scan the list and find the correct message, as all other messages on the list are likely to be gibberish. In the voting context, one cannot simply disregard some potential ground truths as nonsensical; we therefore select a ranking that is close to every possible ground truth.

Our model is also reminiscent of the distance rationalizability framework from the social choice literature [28]. In this framework, there is a fixed set of "consensus profiles" that admit an obvious output. Given a profile of votes, one finds the closest consensus profile (according to some metric), and returns the obvious output for that profile. Our model closely resembles the case where the consensus profiles are strongly unanimous, i.e., they consist of repetitions of a single ranking (which is also the ideal output). The key difference in our model is that instead of focusing solely on the closest ranking (strongly unanimous profile), we need to consider all rankings up to an average distance of t from the given profile — as they are all plausible ground truths — and return a single ranking that is at distance at most k from all such rankings.

A bit further afield, Procaccia et al. [31] study a probabilistic noisy voting setting, and quantify the *robustness* of voting rules to random errors. Their results focus on the probability that the outcome would change, under a random transposition of two adjacent alternatives in a single vote from a submitted profile, in the *worst-case* over profiles. Their work is different from ours in many ways, but perhaps most importantly, they are interested in how frequently common voting rules make mistakes, whereas we are interested in the guarantees of *optimal* voting rules that *avoid* mistakes.

2. Preliminaries

Let A be the set of alternatives, and $|A| = m$. Let $\mathcal{L}(A)$ be the set of rankings over A. A vote σ is a ranking in $\mathcal{L}(A)$, and a profile $\pi \in \mathcal{L}(A)^n$ is a collection of n rankings. A *voting rule* $f : \mathcal{L}(A)^n \to \mathcal{L}(A)$ maps every profile to a ranking.[1]

We assume that there exists an underlying ground truth ranking $\sigma^* \in \mathcal{L}(A)$ of the alternatives, and the votes are noisy estimates of σ^*. We use a distance metric d over $\mathcal{L}(A)$ to measure errors; the error of a vote σ with respect to σ^* is $d(\sigma, \sigma^*)$, and the average error of a profile π with respect to σ^* is $d(\pi, \sigma^*) = (1/n) \cdot \sum_{\sigma \in \pi} d(\sigma, \sigma^*)$. We consider four popular distance metrics over rankings in this paper.

- *The Kendall tau (KT) distance*, denoted d_{KT}, measures the number of pairs of alternatives over which two rankings disagree. Equivalently, it is also the minimum number of swaps of adjacent alternatives required to convert one ranking into another.
- *The (Spearman's) Footrule (FR) distance*, denoted d_{FR}, measures the total displacement of all alternatives between two rankings, i.e., the sum of the absolute differences between their positions in two rankings.
- *The Maximum Displacement (MD) distance*, denoted d_{MD}, measures the maximum of the displacements of all alternatives between two rankings.
- *The Cayley (CY) distance*, denoted d_{CY}, measures the minimum number of swaps (not necessarily of adjacent alternatives) required to convert one ranking into another.

All four metrics described above are *neutral*: A distance metric is called neutral if the distance between two rankings is independent of the labels of the alternatives; in other words, choosing a relabeling of the alternatives and applying it to two rankings keeps the distance between them invariant.

3. Worst-case optimal rules

Suppose we are given a profile π of n noisy rankings that are estimates of an underlying true ranking σ^*. In the absence of any additional information, any ranking could potentially be the true ranking. However, because essentially all crowdsourcing methods draw their power from the often-observed fact that individual opinions are accurate on average, we can plausibly assume that while some agents may make many mistakes, the average error is fairly small. An upper bound on the average error may be inferred by observing the collected votes, or from historical data (but see the next section for the case where this bound is inaccurate).

Formally, suppose we are guaranteed that the average distance between the votes in π and the ground truth σ^* is *at most t* according to a metric d, i.e., $d(\pi, \sigma^*) \le t$. With this guarantee, the set of possible ground truths is given by the "ball" of radius t around π.

$$\mathcal{B}_t^d(\pi) = \{\sigma \in \mathcal{L}(A) \mid d(\pi, \sigma) \le t\}.$$

Note that we have $\sigma^* \in \mathcal{B}_t^d(\pi)$ given our assumption; hence, $\mathcal{B}_t^d(\pi) \neq \emptyset$. We wish to find a ranking that is as close to the ground truth as possible. Since our approach is worst case in nature, our goal is to find the ranking that minimizes the maximum distance from the possible ground truths in $\mathcal{B}_t^d(\pi)$. For a set of rankings $S \subseteq \mathcal{L}(A)$, let its *minimax ranking*, denoted $\textsc{MiniMax}^d(S)$, be defined as follows.[2]

$$\textsc{MiniMax}^d(S) = \underset{\sigma \in \mathcal{L}(A)}{\arg\min} \max_{\sigma' \in S} d(\sigma, \sigma').$$

[1] They are known as *social welfare functions*, which differ from *social choice functions* that choose a single winning alternative.

[2] We use $\textsc{MiniMax}^d(S)$ to denote a single ranking. Ties among multiple minimizers can be broken arbitrarily; our results are independent of the tie-breaking scheme.

Table 1
Application of the optimal voting rules on π.

Voting rule	Possible ground truths $\mathcal{B}_t^d(\pi)$	Output ranking
$\mathrm{OPT}^{d_{KT}}(1.5,\pi)$, $\mathrm{OPT}^{d_{CY}}(1,\pi)$	$\left\{\begin{array}{c} a \succ b \succ c, \\ a \succ c \succ b, \\ b \succ a \succ c \end{array}\right\}$	$a \succ b \succ c$
$\mathrm{OPT}^{d_{FR}}(2,\pi)$, $\mathrm{OPT}^{d_{MD}}(1,\pi)$	$\left\{\begin{array}{c} a \succ b \succ c, \\ a \succ c \succ b \end{array}\right\}$	$\left\{\begin{array}{c} a \succ b \succ c, \\ a \succ c \succ b \end{array}\right\}$

Let the *minimax distance* of S, denoted $k^d(S)$, be the maximum distance of $\textsc{MiniMax}^d(S)$ from the rankings in S according to d. Thus, given a profile π and the guarantee that $d(\pi,\sigma^*) \le t$, the worst-case optimal voting rule OPT^d returns the minimax ranking of the set of possible ground truths $\mathcal{B}_t^d(\pi)$. That is, for all profiles $\pi \in \mathcal{L}(A)^n$ and $t > 0$,

$$\mathrm{OPT}^d(t,\pi) = \textsc{MiniMax}^d\left(\mathcal{B}_t^d(\pi)\right).$$

Furthermore, the output ranking is guaranteed to be at distance at most $k^d(\mathcal{B}_t^d(\pi))$ from the ground truth. We overload notation, and denote $k^d(t,\pi) = k^d(\mathcal{B}_t^d(\pi))$, and

$$k^d(t) = \max_{\pi \in \mathcal{L}(A)^n} k^d(t,\pi).$$

While k^d is explicitly a function of t, it is also implicitly a function of n. Hereinafter, we omit the superscript d whenever the metric is clear from context. Let us illustrate our terminology with a simple example.

Example 1. Let $A = \{a,b,c\}$. We are given profile π consisting of 5 votes: $\pi = \{2 \times (a \succ b \succ c),\ a \succ c \succ b,\ b \succ a \succ c,\ c \succ a \succ b\}$.

The maximum distances between rankings in $\mathcal{L}(A)$ allowed by d_{KT}, d_{FR}, d_{MD}, and d_{CY} are 3, 4, 2, and 2, respectively; let us assume that the average error limit is half the maximum distance for all four metrics.[3]

Consider the Kendall tau distance with $t = 1.5$. The average distances of all 6 rankings from π are given below.

$$\begin{array}{l|l} d_{KT}(\pi, a \succ b \succ c) = 0.8 & d_{KT}(\pi, a \succ c \succ b) = 1.0 \\ d_{KT}(\pi, b \succ a \succ c) = 1.4 & d_{KT}(\pi, b \succ c \succ a) = 2.0 \\ d_{KT}(\pi, c \succ a \succ b) = 1.6 & d_{KT}(\pi, c \succ b \succ a) = 2.2 \end{array}$$

Thus, the set of possible ground truths is $\mathcal{B}_{1.5}^{d_{KT}}(\pi) = \{a \succ b \succ c,\ a \succ c \succ b,\ b \succ a \succ c\}$. This set has a unique minimax ranking $\mathrm{OPT}^{d_{KT}}(1.5,\pi) = a \succ b \succ c$, which gives $k^{d_{KT}}(1.5,\pi) = 1$. Table 1 lists the sets of possible ground truths and their minimax rankings[4] under different distance metrics.

Note that even with identical (scaled) error bounds, different distance metrics lead to different sets of possible ground truths as well as different optimal rankings. This demonstrates that the choice of the distance metric is important.

3.1. Upper bound

Given a distance metric d, a profile π, and that $d(\pi,\sigma^*) \le t$, we can bound $k(t,\pi)$ using the *diameter* of the set of possible ground truths $\mathcal{B}_t(\pi)$. For a set of rankings $S \subseteq \mathcal{L}(A)$, denote its diameter by $\mathcal{D}(S) = \max_{\sigma,\sigma' \in S} d(\sigma,\sigma')$.

Lemma 1. $\frac{1}{2} \cdot \mathcal{D}(\mathcal{B}_t(\pi)) \le k(t,\pi) \le \mathcal{D}(\mathcal{B}_t(\pi)) \le 2t$.

Proof. Let $\widehat{\sigma} = \textsc{MiniMax}(\mathcal{B}_t(\pi))$. For rankings $\sigma,\sigma' \in \mathcal{B}_t(\pi)$, we have $d(\sigma,\widehat{\sigma}), d(\sigma',\widehat{\sigma}) \le k(t,\pi)$ by definition of $\widehat{\sigma}$. By the triangle inequality, $d(\sigma,\sigma') \le 2k(t,\pi)$ for all $\sigma,\sigma' \in \mathcal{B}_t(\pi)$. Thus, $\mathcal{D}(\mathcal{B}_t(\pi)) \le 2k(t,\pi)$.

Next, the maximum distance of $\sigma \in \mathcal{B}_t(\pi)$ from all rankings in $\mathcal{B}_t(\pi)$ is at most $\mathcal{D}(\mathcal{B}_t(\pi))$. Hence, the minimax distance $k(t,\pi) = k(\mathcal{B}_t(\pi))$ cannot be greater than $\mathcal{D}(\mathcal{B}_t(\pi))$.

Finally, let $\pi = \{\sigma_1,\dots,\sigma_n\}$. For rankings $\sigma,\sigma' \in \mathcal{B}_t(\pi)$, the triangle inequality implies $d(\sigma,\sigma') \le d(\sigma,\sigma_i) + d(\sigma_i,\sigma')$ for every $i \in \{1,\dots,n\}$. Averaging over these inequalities, we get $d(\sigma,\sigma') \le t + t = 2t$, for all $\sigma,\sigma' \in \mathcal{B}_t(\pi)$. Thus, we have $\mathcal{D}(\mathcal{B}_t(\pi)) \le 2t$, as required. □

Lemma 1 implies that $k(t) = \max_{\pi \in \mathcal{L}(A)^n} k(t,\pi) \le 2t$ for all distance metrics and $t > 0$. In words:

[3] Scaling by the maximum distance is not a good way of comparing distance metrics; we do so for the sake of illustration only.
[4] Multiple rankings indicate a tie that can be broken arbitrarily.

Theorem 1. *Given n noisy rankings at an average distance of at most t from an unknown true ranking σ^* according to a distance metric d, it is always possible to find a ranking at distance at most $2t$ from σ^* according to d.*

Importantly, the bound of Theorem 1 is independent of the number of votes n. Most statistical models of social choice restrict profiles in two ways: i) the average error should be low because the probability of generating high-error votes is typically low, and ii) the errors should be distributed almost evenly (in different directions from the ground truth), which is why aggregating the votes works well. These assumptions are mainly helpful when n is large, that is, performance may be poor for small n (see, e.g., [12]). In contrast, our model restricts profiles only by making the first assumption (explicitly), allowing voting rules to perform well as long as the votes are accurate on average, independently of the number of votes n.

We also remark that Theorem 1 admits a simple proof, but the bound is nontrivial: while the average error of the profile is at most t (hence, the profile contains a ranking with error at most t), it is generally impossible to pinpoint a single ranking within the profile that has error at most $2t$ with respect to the ground truth in the worst-case (i.e., with respect to every possible ground truth in $\mathcal{B}_t(\pi)$). That said, it can be shown that there exists a ranking in the profile that always has distance at most $3t$ from the ground truth. Further, one can pick such a ranking in polynomial time, which stands in sharp contrast to the usual hardness of finding the optimal ranking (see the discussion on the computational complexity of our approach in Section 6).

Theorem 2. *Given n noisy rankings at an average distance of at most t from an unknown true ranking σ^* according to a distance metric d, it is always possible to pick, in polynomial time, one of the n given rankings that has distance at most $3t$ from σ^* according to d.*

Proof. Consider a profile π consisting of n rankings such that $d(\sigma^*, \pi) \le t$. Let $x = \min_{\sigma \in \mathcal{L}(A)} d(\sigma, \pi)$ be the minimum distance any ranking has from the profile. Then, $x \le d(\sigma^*, \pi) \le t$. Let $\widehat{\sigma} = \arg\min_{\sigma \in \pi} d(\sigma, \pi)$ be the ranking in π which minimizes the distance from π among all rankings in π. An easy-to-verify folklore theorem says that $d(\widehat{\sigma}, \pi) \le 2x$. To see this, assume that ranking τ has the minimum distance from the profile (i.e., $d(\tau, \pi) = x$). Now, the average distance of all rankings in π from π is

$$\begin{aligned}\frac{1}{n}\sum_{\sigma \in \pi} d(\sigma, \pi) = \frac{1}{n^2}\sum_{\sigma \in \pi}\sum_{\sigma' \in \pi} d(\sigma, \sigma') &\le \frac{1}{n^2}\sum_{\sigma \in \pi}\sum_{\sigma' \in \pi}(d(\tau, \sigma) + d(\tau, \sigma')) \\ &= \frac{2}{n}\sum_{\sigma \in \pi} d(\tau, \sigma) = 2x \le 2t,\end{aligned}$$

where the second transition uses the triangle inequality. Now, $\widehat{\sigma}$ has the smallest distance from π among all rankings in π, which cannot be greater than the average distance $(1/n)\sum_{\sigma \in \pi} d(\sigma, \pi)$. Hence, $d(\widehat{\sigma}, \pi) \le 2t$. Finally,

$$d(\widehat{\sigma}, \sigma^*) \le \frac{1}{n}\sum_{\sigma \in \pi}\left(d(\widehat{\sigma}, \sigma) + d(\sigma, \sigma^*)\right) \le 2t + t = 3t,$$

where the first transition uses the triangle inequality and the second transition uses the fact that $d(\widehat{\sigma}, \pi) \le 2t$ and $d(\pi, \sigma^*) \le t$. It is easy to see that $\widehat{\sigma}$ can be computed in $O(n^2)$ time. □

3.2. Lower bounds

The upper bound of $2t$ (Theorem 1) is intuitively loose — we cannot expect it to be tight for every distance metric. However, we can complement it with a lower bound of (roughly speaking) $t/2$ for all distance metrics. Formally, let $d^{\downarrow}(r)$ denote the greatest feasible distance under distance metric d that is less than or equal to r. Next, we prove a lower bound of $d^{\downarrow}(t)/2$.

Theorem 3. *For a distance metric d, $k(t) \ge d^{\downarrow}(t)/2$.*

Proof. If $d^{\downarrow}(t) = 0$, then the result trivially holds. Assume $d^{\downarrow}(t) > 0$. Let σ and σ' be two rankings at distance $d^{\downarrow}(t)$. Consider profile π consisting of only a single instance of ranking σ. Then, $\sigma' \in \mathcal{B}_t(\pi)$. Hence, $\mathcal{D}(\mathcal{B}_t(\pi)) \ge d^{\downarrow}(t)$. Now, it follows from Lemma 1 that $k(t) \ge \mathcal{D}(\mathcal{B}_t(\pi))/2 \ge d^{\downarrow}(t)/2$. □

Recall that Theorem 1 shows that $k(t) \le 2t$. However, $k(t)$ is the minimax distance under some profile, and hence must be a feasible distance under d. Thus, Theorem 1 actually implies a possibly better upper bound of $d^{\downarrow}(2t)$. Together with Theorem 3, this implies $d^{\downarrow}(t)/2 \le k(t) \le d^{\downarrow}(2t)$. Next, we show that imposing a mild assumption on the distance metric allows us to improve the lower bound by a factor of 2, thus reducing the gap between the lower and upper bounds.

Theorem 4. *For a neutral distance metric d, $k(t) \ge d^{\downarrow}(t)$.*

Proof. For a ranking $\sigma \in \mathcal{L}(A)$ and $r \geq 0$, let $\mathcal{B}_r(\sigma)$ denote the set of rankings at distance at most r from σ. Neutrality of the distance metric d implies $|\mathcal{B}_r(\sigma)| = |\mathcal{B}_r(\sigma')|$ for all $\sigma, \sigma' \in \mathcal{L}(A)$ and $r \geq 0$. In particular, $d^{\downarrow}(t)$ being a feasible distance under d implies that for every $\sigma \in \mathcal{L}(A)$, there exists some ranking at distance exactly $d^{\downarrow}(t)$ from σ.

Fix $\sigma \in \mathcal{L}(A)$. Consider the profile π consisting of n instances of σ. It holds that $\mathcal{B}_t(\pi) = \mathcal{B}_t(\sigma)$. We want to show that the minimax distance $k(\mathcal{B}_t(\sigma)) \geq d^{\downarrow}(t)$. Suppose for contradiction that there exists some $\sigma' \in \mathcal{L}(A)$ such that all rankings in $\mathcal{B}_t(\sigma)$ are at distance at most t' from σ', i.e., $\mathcal{B}_t(\sigma) \subseteq \mathcal{B}_{t'}(\sigma')$, with $t' < d^{\downarrow}(t)$. Since there exists some ranking at distance $d^{\downarrow}(t) > t'$ from σ', we have $\mathcal{B}_t(\sigma) \subseteq \mathcal{B}_{t'}(\sigma') \subsetneq \mathcal{B}_t(\sigma')$, which is a contradiction because $|\mathcal{B}_t(\sigma)| = |\mathcal{B}_t(\sigma')|$. Therefore, $k(t) \geq k(t, \pi) \geq d^{\downarrow}(t)$. $\square$

The bound of Theorem 4 holds for all $n, m > 0$ and all $t \in [0, D]$, where D is the maximum possible distance under d. It can be checked easily that the bound is tight given the neutrality assumption, which is an extremely mild — and in fact, a highly desirable — assumption for distance metrics over rankings.

Theorem 4 improves the bounds on $k(t)$ to $d^{\downarrow}(t) \leq k(t) \leq d^{\downarrow}(2t)$ for a variety of distance metrics d. However, for the four special distance metrics considered in this paper, the next result, which is our main theoretical result, closes this gap by establishing a tight lower bound of $d^{\downarrow}(2t)$, for a wide range of values of n and t.

Theorem 5. *If $d \in \{d_{KT}, d_{FR}, d_{MD}, d_{CY}\}$, and the maximum distance allowed by the metric is $D \in \Theta(m^{\alpha})$, then there exists $T \in \Theta(m^{\alpha})$ such that:*

1. *For all $t \leq T$ and even n, we have $k(t) \geq d^{\downarrow}(2t)$.*
2. *For all $L \geq 2$, $t \leq T$ with $\{2t\} \in (1/L, 1 - 1/L)$, and odd $n \geq \Theta(L \cdot D)$, we have $k(t) \geq d^{\downarrow}(2t)$. Here, $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of $x \in \mathbb{R}$.*

The impossibility result of Theorem 5 is weaker for odd values of n (in particular, covering more values of t requires larger n), which is reminiscent of the fact that repetition (error-correcting) codes achieve greater efficiency with an odd number of repetitions; this is not merely a coincidence. Indeed, an extra repetition allows differentiating between tied possibilities for the ground truth; likewise, an extra vote in the profile prevents us from constructing a symmetric profile that admits a diverse set of possible ground truths.

Proof of Theorem 5. We denote $\{1, \ldots, r\}$ by $[r]$ in this proof. We use $\sigma(a)$ to denote the rank (position) of alternative a in ranking σ. First, we prove the case of even n for all four distance metrics. We later provide a generic argument to prove the case of large odd n. First, we need a simple observation.

Observation 1. *If $\binom{r}{2} \leq \lfloor 2t \rfloor$ and $t \geq 0.5$, then $r \leq 4\sqrt{t}$.*

Proof. Note that $(r-1)^2 \leq r \cdot (r-1) \leq 2 \cdot \lfloor 2t \rfloor \leq 4t$. Hence, $r \leq 2\sqrt{t} + 1$. We also have $t \geq 0.5$, i.e., $1 \leq 2t$. This implies $1 \leq \sqrt{2t}$. Thus, we have $r \leq 2\sqrt{t} + \sqrt{2t} = (2 + \sqrt{2})\sqrt{t} \leq 4\sqrt{t}$. $\square$

The Kendall tau distance: Let d be the Kendall tau distance; thus, $D = \binom{m}{2}$ and $\alpha = 2$. Let n be even. For a ranking $\tau \in \mathcal{L}(A)$, let τ_{rev} be its reverse. Assume $t = (1/2) \cdot \binom{m}{2}$, and fix a ranking $\sigma \in \mathcal{L}(A)$. Every ranking must agree with exactly one of σ and σ_{rev} on a given pair of alternatives. Hence, every $\rho \in \mathcal{L}(A)$ satisfies $d(\rho, \sigma) + d(\rho, \sigma_{\text{rev}}) = \binom{m}{2}$. Consider the profile π consisting of $n/2$ instances of σ and $n/2$ instances of σ_{rev}. Then, the average distance of every ranking from rankings in π would be exactly t, i.e., $\mathcal{B}_t(\pi) = \mathcal{L}(A)$. It is easy to check that $k(\mathcal{L}(A)) = \binom{m}{2} = 2t = d^{\downarrow}(2t)$ because every ranking has its reverse ranking in $\mathcal{L}(A)$ at distance exactly $2t$.

Now, let us extend the proof to $t \leq (m/12)^2$. If $t < 0.5$, then $d^{\downarrow}_{KT}(2t) = 0$, which is a trivial lower bound. Hence, assume $t \geq 0.5$. Thus, $d^{\downarrow}(2t) = \lfloor 2t \rfloor$. We use Fermat's Polygonal Number Theorem (see, e.g., [22]). A special case of this remarkable theorem states that every natural number can be expressed as the sum of at most three "triangular" numbers, i.e., numbers of the form $\binom{k}{2}$. Let $\lfloor 2t \rfloor = \sum_{i=1}^{3} \binom{m_i}{2}$. From Observation 1, it follows that $0 \leq m_i \leq 4\sqrt{t}$ for all $i \in \{1, 2, 3\}$. Hence, $\sum_{i=1}^{3} m_i \leq 12\sqrt{t} \leq m$.

Partition the set of alternatives A into four disjoint groups A_1, A_2, A_3, and A_4 such that $|A_i| = m_i$ for $i \in \{1, 2, 3\}$, and $|A_4| = m - \sum_{i=1}^{3} m_i$. Let σ^{A_4} be an arbitrary ranking of the alternatives in A_4; consider the partial order $\mathcal{P}_A = A_1 \succ A_2 \succ A_3 \succ \sigma^{A_4}$ over alternatives in A. Note that a ranking ρ is an extension of $\mathcal{P}_A$ iff it ranks all alternatives in A_i before any alternative in A_{i+1} for $i \in \{1, 2, 3\}$, and ranks alternatives in A_4 according to σ^{A_4}. Choose arbitrary $\sigma^{A_i} \in \mathcal{L}(A_i)$ for $i \in \{1, 2, 3\}$ and define

$$\sigma = \sigma^{A_1} \succ \sigma^{A_2} \succ \sigma^{A_3} \succ \sigma^{A_4},$$
$$\sigma' = \sigma^{A_1}_{\text{rev}} \succ \sigma^{A_2}_{\text{rev}} \succ \sigma^{A_3}_{\text{rev}} \succ \sigma^{A_4}.$$

Note that both σ and σ' are extensions of $\mathcal{P}_A$. Once again, take the profile π consisting of $n/2$ instances of σ and $n/2$ instances of σ'. It is easy to check that a ranking disagrees with exactly one of σ and σ' on every pair of alternatives that belong to the same group in $\{A_1, A_2, A_3\}$. Hence, every ranking $\rho \in \mathcal{L}(A)$ satisfies

$$d(\rho,\sigma) + d(\rho,\sigma') \geq \sum_{i=1}^{3} \binom{m_i}{2} = \lfloor 2t \rfloor . \tag{1}$$

Clearly an equality is achieved in Equation (1) if and only if ρ is an extension of $\mathcal{P}_A$. Thus, every extension of $\mathcal{P}_A$ has an average distance of $\lfloor 2t \rfloor /2 \leq t$ from π. Every ranking ρ that is not an extension of $\mathcal{P}_A$ achieves a strict inequality in Equation (1); thus, $d(\rho,\pi) \geq (\lfloor 2t \rfloor + 1)/2 > t$. Hence, $\mathcal{B}_t(\pi)$ is the set of extensions of $\mathcal{P}_A$.

Given a ranking $\rho \in \mathcal{L}(A)$, consider the ranking in $\mathcal{B}_t(\pi)$ that reverses the partial orders over A_1, A_2, and A_3 induced by ρ. The distance of this ranking from ρ would be at least $\sum_{i=1}^{3} \binom{m_i}{2} = \lfloor 2t \rfloor$, implying $k(\mathcal{B}_t(\pi)) \geq \lfloor 2t \rfloor$. (In fact, it can be checked that $k(\mathcal{B}_t(\pi)) = \mathcal{D}(\mathcal{B}_t(\pi)) = \lfloor 2t \rfloor$.)

We now proceed to prove the case of an even number of agents for the other three distance metrics. First, if M is the minimum distance between two distinct rankings under a distance metric d and $t < M/2$, then we have $d^{\downarrow}(2t) = 0$, which is a trivial lower bound. Hence, we assume $t \geq M/2$.

The footrule distance: Let d_{FR} denote the footrule distance; recall that given $\sigma, \sigma' \in \mathcal{L}(A)$, $d_{FR}(\sigma,\sigma') = \sum_{a \in A} |\sigma(a) - \sigma'(a)|$. The proof is along the same lines as the proof for the Kendall tau distance, but uses a few additional clever ideas. It is known that the maximum footrule distance between two rankings over m alternatives is $D = \lfloor m^2/2 \rfloor$, and is achieved by two rankings that are reverse of each other [17]. Hence, we have $\alpha = 2$; thus, we wish to find $T \in \Theta(m^2)$ for which the claim will hold. Formally writing the distance between a ranking and its reverse, we get

$$d_{FR}(\sigma, \sigma_{\text{rev}}) = \sum_{i=1}^{m} |m+1-2i| = \left\lfloor \frac{m^2}{2} \right\rfloor . \tag{2}$$

Observation 2. *The footrule distance between two rankings is always an even integer.*

Proof. Take rankings $\sigma, \tau \in \mathcal{L}(A)$. Note that $d_{FR}(\sigma,\tau) = \sum_{a \in A} |\sigma(a) - \tau(a)|$. Now, $|\sigma(a) - \tau(a)|$ is odd if and only if the positions of a in σ and τ have different parity. Since the number of odd (as well as even) positions is identical in σ and τ, the number of alternatives that leave an even position in σ to go to an odd position in τ equals the number of alternatives that leave an odd position in σ to go to an even position in τ. Thus, the number of alternatives for which the parity of the position changes is even. Equivalently, the number of odd terms in the sum defining the footrule distance is even. Hence, the footrule distance is an even integer. □

Hence, Equation (2) implies that $d^{\downarrow}_{FR}(2t)$ equals $\lfloor 2t \rfloor$ if $\lfloor 2t \rfloor$ is even, and equals $\lfloor 2t \rfloor - 1$ otherwise. Let $r = d^{\downarrow}_{FR}(2t)$. Hence, r is an even integer. We prove the result for $t \leq (m/8)^2$. In this case, we invoke the 4-gonal special case of Fermat's Polygonal Number Theorem (instead of the 3-gonal case invoked in the proof for the Kendall tau distance): *Every positive integer can be written as the sum of at most four squares.* Let $r/2 = m_1^2 + m_2^2 + m_3^2 + m_4^2$. Hence,

$$r = \frac{(2m_1)^2}{2} + \frac{(2m_2)^2}{2} + \frac{(2m_3)^2}{2} + \frac{(2m_4)^2}{2} . \tag{3}$$

It is easy to check that $m_i \leq \sqrt{r/2}$ for $i \in [4]$. Thus, $\sum_{i=1}^{4} 2m_i \leq 8\sqrt{r/2} \leq 8\sqrt{t} \leq m$. Let us partition the set of alternatives A into $\{A_i\}_{i \in [5]}$ such that $|A_i| = 2m_i$ for $i \in [4]$ and $|A_5| = m_5 = m - \sum_{i=1}^{4} 2m_i$.

Fix $\sigma^{A_5} \in \mathcal{L}(A_5)$ and consider the partial order $\mathcal{P}_A = A_1 \succ A_2 \succ A_3 \succ A_4 \succ \sigma^{A_5}$. Choose arbitrary $\sigma^{A_i} \in \mathcal{L}(A_i)$ for $i \in [4]$, and let

$$\sigma = (\sigma^{A_1} \succ \sigma^{A_2} \succ \sigma^{A_3} \succ \sigma^{A_4} \succ \sigma^{A_5}),$$
$$\sigma' = (\sigma^{A_1}_{\text{rev}} \succ \sigma^{A_2}_{\text{rev}} \succ \sigma^{A_3}_{\text{rev}} \succ \sigma^{A_4}_{\text{rev}} \succ \sigma^{A_5}).$$

Note that both σ and σ' are extensions of $\mathcal{P}_A$. Consider the profile π consisting of $n/2$ instances of σ and σ' each. Unlike the Kendall tau distance, $\mathcal{B}_t(\pi)$ is not the set of extensions of $\mathcal{P}_A$. Still, we show that it satisfies $k(\mathcal{B}_t(\pi)) = \mathcal{D}(\mathcal{B}_t(\pi)) = d^{\downarrow}_{FR}(2t) = r$.

Denote by a_i^j the alternative ranked j in σ^{A_i}. Take a ranking $\rho \in \mathcal{L}(A)$. Consider $d_{FR}(\rho,\sigma) + d_{FR}(\rho,\sigma')$. We have the following inequalities regarding the sum of displacement of different alternatives between ρ and σ, and between ρ and σ'. For $i \in [4]$ and $j \in [2m_i]$,

$$\left|\rho(a_i^j) - \sigma(a_i^j)\right| + \left|\rho(a_i^j) - \sigma'(a_i^j)\right| \geq \left|\sigma(a_i^j) - \sigma'(a_i^j)\right| = |j - (2m_i + 1 - j)|. \tag{4}$$

Summing all the inequalities, we get

$$d_{FR}(\rho,\sigma)+d_{FR}(\rho,\sigma')\geq\sum_{i=1}^{4}\sum_{j=1}^{2m_i}|2j-2m_i-1|=\sum_{i=1}^{4}\frac{(2m_i)^2}{2}=r, \tag{5}$$

where the second transition follows from Equation (2), and the third transition follows from Equation (3).

First, we show that $\rho\in\mathcal{B}_t(\pi)$ only if equality in Equation (5) holds. To see why, note that the footrule distance is always even and $r=d^{\downarrow}_{FR}(2t)\geq\lfloor 2t\rfloor-1$. Hence, if equality is not achieved, then $d_{FR}(\rho,\sigma)+d_{FR}(\rho,\sigma')\geq r+2\geq\lfloor 2t\rfloor-1+1>2t$. Hence, the average distance of ρ from votes in π would be greater than t.

On the contrary, if equality is indeed achieved in Equation (5), then the average distance of ρ from votes in π is $r/2\leq t$. Hence, we have established that $\mathcal{B}_t(\pi)$ is the set of rankings ρ for which equality is achieved in Equation (5).

For ρ to achieve equality in Equation (5), it must achieve equality in Equation (4) for every $i\in[4]$ and $j\in[2m_i]$, and it must agree with both σ and σ' on the positions of alternatives in A_5 (i.e., σ^{A_5} must be a suffix of ρ). For the former to hold, the position of a_i^j in ρ must be between $\sigma(a_i^j)$ and $\sigma'(a_i^j)=\sigma(a_i^{2m_i+1-j})$ (both inclusive), for every $i\in[4]$ and $j\in[2m_i]$.

We claim that the set of rankings satisfying these conditions are characterized as follows.

$$\begin{aligned}\mathcal{B}_t(\pi)=\Big\{\rho\in\mathcal{L}(A)\ \Big|\ &\{\rho(a_i^j),\rho(a_i^{2m_i+1-j})\}=\{\sigma(a_i^j),\sigma(a_i^{2m_i+1-j})\}\\ &\text{for } i\in[4],\ j\in[2m_i],\ \text{and}\\ &\rho(a_5^j)=\sigma(a_5^j)=\sigma'(a_5^j)\text{ for } j\in[m_5]\Big\}.\end{aligned} \tag{6}$$

Note that instead of $\rho(a_i^j)$ and $\rho(a_i^{2m_i+1-j})$ both being in the interval $[\sigma(a_i^j),\sigma(a_i^{2m_i+1-j})]$, we are claiming that they must be the two endpoints. First, consider the middle alternatives in each A_i ($i\in[4]$), namely $a_i^{m_i}$ and $a_i^{m_i+1}$. Both must be placed between $\sigma(a_i^{m_i})=\sigma'(a_i^{m_i+1})$ and $\sigma(a_i^{m_i+1})=\sigma'(a_i^{m_i})$; but these two numbers differ by exactly 1. Hence,

$$\left\{\rho(a_i^{m_i}),\rho(a_i^{m_i+1})\right\}=\left\{\sigma(a_i^{m_i}),\sigma(a_i^{m_i+1})\right\}.$$

Consider the two adjacent alternatives, namely $a_i^{m_i-1}$ and $a_i^{m_i+2}$. Given that the middle alternatives $a_i^{m_i}$ and $a_i^{m_i+1}$ occupy their respective positions in σ or σ', the only positions available to ρ for placing the two adjacent alternatives are the endpoints of their common feasible interval $[\sigma(a_i^{m_i-1}),\sigma(a_i^{m_i+2})]$. Continuing this argument, each pair of alternatives $(a_i^j,a_i^{2m_i+1-j})$ must occupy the two positions $\{\sigma(a_i^j),\sigma(a_i^{2m_i+1-j})\}$ for every $i\in[4]$ and $j\in[m_i]$.

That is, ρ can either keep the alternatives a_i^j and $a_i^{2m_i+1-j}$ as they are in σ, or place them according to σ' (equivalently, swapping them in σ) for every $i\in[4]$ and $j\in[2m_i]$. Note that these choices are independent of each other. We established that a ranking ρ is in $\mathcal{B}_t(\pi)$ only if it is obtained in this manner and has σ^{A_5} as its suffix.

Further, it can be seen that each of these choices (keeping or swapping the pair in σ) maintain $d_{FR}(\rho,\sigma)+d_{FR}(\rho,\sigma')$ invariant. Hence, all such rankings ρ satisfy $d_{FR}(\rho,\sigma)+d_{FR}(\rho,\sigma')=r$, and thus belong to $\mathcal{B}_t(\pi)$. This reaffirms our original claim that $\mathcal{B}_t(\pi)$ is given by Equation (6).

In summary, all rankings in $\mathcal{B}_t(\pi)$ can be obtained by taking σ, and arbitrarily choosing whether to swap the pair of alternatives a_i^j and $a_i^{2m_i+1-j}$ for each $i\in[4]$ and $j\in[2m_i]$.

Note that $\sigma,\sigma'\in\mathcal{B}_t(\pi)$ and $d_{FR}(\sigma,\sigma')=r$ (this distance is given by the summation in Equation (5)). Hence, $\mathcal{D}(\mathcal{B}_t(\pi))\geq r$. Now, we prove that its minimax distance is at least r as well. Take a ranking $\rho\in\mathcal{L}(A)$. We need to show that there exists some $\tau\in\mathcal{B}_t(\pi)$ such that $d_{FR}(\rho,\tau)\geq r$.

Consider alternatives a_i^j and $a_i^{2m_i+1-j}$ for $i\in[4]$ and $j\in[2m_i]$. We know that τ must satisfy $\{\tau(a_i^j),\tau(a_i^{2m_i+1-j})\}=\{\sigma(a_i^j),\sigma(a_i^{2m_i+1-j})\}$ in order to belong to $\mathcal{B}_t(\pi)$. This allows two possible ways for placing the pair of alternatives. Let τ pick the optimal positions that maximize

$$\tau_{i,j}(\rho)=|\tau(a_i^j)-\rho(a_i^j)|+|\tau(a_i^{2m_i+1-j})-\rho(a_i^{2m_i+1-j})|.$$

That is, $\tau_{i,j}(\rho)$ should equal $M_{i,j}(\rho)$, which we define as

$$\begin{aligned}\max\Big\{&|\sigma(a_i^j)-\rho(a_i^j)|+|\sigma(a_i^{2m_i+1-j})-\rho(a_i^{2m_i+1-j})|,\\ &|\sigma(a_i^{2m_i+1-j})-\rho(a_i^j)|+|\sigma(a_i^j)-\rho(a_i^{2m_i+1-j})|\Big\}.\end{aligned}$$

Note that the choice for each pair of alternatives $(a_i^j,a_i^{2m_i+1-j})$ can be made independently of every other pair. Further, making the optimal choice for each pair guarantees that $d_{FR}(\rho,\tau)$ is at least

$$\sum_{i=1}^{4}\sum_{j=1}^{2m_i}\tau_{i,j}(\rho) = \sum_{i=1}^{4}\sum_{j=1}^{2m_i} M_{i,j}(\rho),$$

which we will now show to be at least r.

Algorithm 1 describes how to find the optimal ranking $\tau \in \mathcal{B}_t(\pi)$ mentioned above, which satisfies $\tau_{i,j}(\rho) = M_{i,j}(\rho)$ for every $i \in [4]$ and $j \in [2m_i]$. It starts with an arbitrary $\tau \in \mathcal{B}_t(\pi)$, and swaps every sub-optimally placed pair $(a_i^j, a_i^{2m_i+1-j})$ for $i \in [4]$ and $j \in [2m_i]$. In the algorithm, $\tau_{a\leftrightarrow b}$ denotes the ranking obtained by swapping alternatives a and b in τ.

ALGORITHM 1: Finds a ranking in $\mathcal{B}_t(\pi)$ at a footrule distance of at least $\lfloor 2t \rfloor$ from any given ranking.

Data: Ranking $\rho \in \mathcal{L}(A)$
Result: Ranking $\tau \in \mathcal{B}_t(\pi)$ such that $d_{FR}(\tau, \rho) \geq \lfloor 2t \rfloor$
$\tau \leftarrow$ an arbitrary ranking from $\mathcal{B}_t(\pi)$;
for $i \in [4]$ **do**
  for $j \in [2m_i]$ **do**
    $d_i^j \leftarrow |\rho(a_i^j) - \tau(a_i^j)|$;
    $d_i^{2m_i+1-j} \leftarrow |\rho(a_i^{2m_i+1-j}) - \tau(a_i^{2m_i+1-j})|$;
    if $d_i^j + d_i^{2m_i+1-j} < M_{i,j}(\rho)$ **then**
      $\tau \leftarrow \tau_{a_i^j \leftrightarrow a_i^{2m_i+1-j}}$;
    end
  end
end
return τ;

Finally, we show that $d_{FR}(\rho, \tau) \geq r$. First, we establish the following lower bound on $M_{i,j}(\rho)$.

$$\begin{aligned} M_{i,j}(\rho) &\geq \frac{1}{2}\Big(|\sigma(a_i^j) - \rho(a_i^j)| + |\sigma(a_i^{2m_i+1-j}) - \rho(a_i^{2m_i+1-j})| \\ &\qquad + |\sigma(a_i^{2m_i+1-j}) - \rho(a_i^j)| + |\sigma(a_i^j) - \rho(a_i^{2m_i+1-j})|\Big) \\ &\geq |\sigma(a_i^{2m_i+1-j}) - \sigma(a_i^j)| \\ &= |2m_i + 1 - 2j|, \end{aligned}$$

where the first transition holds because the maximum of two terms is at least as much as their average, and the second transition uses the triangle inequality on appropriately paired terms. Now, we have

$$d_{FR}(\tau, \rho) \geq \sum_{i=1}^{4}\sum_{j=1}^{2m_i} M_{i,j}(\rho) \geq \sum_{i=1}^{4}\sum_{j=1}^{2m_i} |2m_i + 1 - 2j| = \sum_{i=1}^{4} \frac{(2m_i)^2}{2} = r,$$

where the third transition holds due to Equation (2), and the fourth transition holds due to Equation (3). Hence, the minimax distance of $\mathcal{B}_t(\pi)$ is at least $r = d_{FR}^{\downarrow}(2t)$, as required.

The Cayley distance: Next, let d_{CY} denote the Cayley distance. Recall that $d_{CY}(\sigma, \tau)$ equals the minimal number of swaps (of possibly non-adjacent alternatives) required in order to transform σ to τ. It is easy to check that the maximum Cayley distance is $D = m - 1$; hence, it has $\alpha = 1$. We prove the result for $t \leq m/4$. Note that $d_{CY}^{\downarrow}(2t) = \lfloor 2t \rfloor$. Define rankings $\sigma, \sigma' \in \mathcal{L}(A)$ as follows.

$$\begin{aligned} \sigma &= (\underbrace{a_1 \succ \ldots \succ a_{2\lfloor 2t \rfloor}} \succ a_{2\lfloor 2t \rfloor + 1} \succ \ldots \succ a_m), \\ \sigma' &= (\underbrace{a_{2\lfloor 2t \rfloor} \succ \ldots \succ a_1} \succ a_{2\lfloor 2t \rfloor + 1} \succ \ldots \succ a_m). \end{aligned}$$

Let profile π consist of $n/2$ instances of σ and σ' each. We claim that $\mathcal{B}_t(\pi)$ has the following structure, which is very similar to the ball for the footrule distance.

$$\begin{aligned} \mathcal{B}_t(\pi) = \Big\{\rho \in \mathcal{L}(A) \;\Big|\; &\{\rho(a_i), \rho(a_{2\lfloor 2t \rfloor + 1 - i})\} = \{i, 2\lfloor 2t \rfloor + 1 - i\} \text{ for } i \in [\lfloor 2t \rfloor], \\ &\text{and } \rho(a_i) = i \text{ for } i > 2\lfloor 2t \rfloor \Big\}. \end{aligned} \tag{7}$$

First, we observe the following simple fact: If rankings τ and ρ mismatch (i.e., place different alternatives) in r different positions, then $d_{CY}(\tau, \rho) \geq r/2$. Indeed, consider the number of swaps required to convert τ into ρ. Since each swap can make τ and ρ consistent in at most two more positions, it would take at least $r/2$ swaps to convert τ into ρ, i.e., $d_{CY}(\tau,\rho) \geq r/2$.

Now, note that σ and σ' mismatch in each of first $2\lfloor 2t \rfloor$ positions. Hence, every ranking $\rho \in \mathcal{L}(A)$ must mismatch with at least one of σ and σ' in each of first $2\lfloor 2t \rfloor$ positions. Together with the previous observation, this implies

$$d_{CY}(\rho,\sigma) + d_{CY}(\rho,\sigma') \geq \lfloor 2t \rfloor. \tag{8}$$

Every ranking ρ that achieves equality in Equation (8) is clearly in $\mathcal{B}_t(\pi)$ because its average distance from the votes in π is $\lfloor 2t \rfloor/2 \leq t$. Further, every ranking ρ that achieves a strict inequality in Equation (8) is outside $\mathcal{B}_t(\pi)$ because its average distance from the votes in π is at least $(\lfloor 2t \rfloor + 1)/2 > t$. Hence, $\mathcal{B}_t(\pi)$ consists of rankings that satisfy $d_{CY}(\rho,\sigma) + d_{CY}(\rho,\sigma') = \lfloor 2t \rfloor$.

Now, any ranking ρ satisfying equality in Equation (8) must be consistent with exactly one of σ and σ' in each of first $2\lfloor 2t \rfloor$ positions, and with both σ and σ' in the later positions. The former condition implies that for every $i \in \lfloor 2t \rfloor$, ρ must place the pair of alternatives $(a_i, a_{2\lfloor 2t \rfloor+1-i})$ in positions i and $2\lfloor 2t \rfloor + 1 - i$, either according to σ or according to σ'. This confirms our claim that $\mathcal{B}_t(\pi)$ is given by Equation (7).

We now show that $k(\mathcal{B}_t(\pi)) \geq \lfloor 2t \rfloor$. Take a ranking $\rho \in \mathcal{L}(A)$. We construct a ranking $\tau \in \mathcal{B}_t(\pi)$ such that τ mismatches with ρ in each of first $2\lfloor 2t \rfloor$ positions. Together with our observation that the Cayley distance is at least half of the number of positional mismatches, this would imply that the minimax distance of $\mathcal{B}_t(\pi)$ is at least $\lfloor 2t \rfloor$, as required.

We construct τ by choosing the placement of the pair of alternatives $(a_i, a_{2\lfloor 2t \rfloor+1-i})$, independently for each $i \in \lfloor 2t \rfloor$, in a way that τ mismatches with ρ in positions i and $2\lfloor 2t \rfloor + 1 - i$ both. Let $\mathbb{I}(X)$ denote the indicator variable that is 1 if statement X holds, and 0 otherwise. Let $r = \mathbb{I}(\rho(a_i) = i) + \mathbb{I}\left(\rho(a_{2\lfloor 2t \rfloor+1-i}) = 2\lfloor 2t \rfloor + 1 - i\right)$. Consider the following three cases.

- $r = 0$: Set $\tau(a_i) = i$ and $\tau(a_{2\lfloor 2t \rfloor+1-i}) = 2\lfloor 2t \rfloor + 1 - i$.
- $r = 1$: Without loss of generality, assume $\rho(a_i) = i$. Set $\tau(a_i) = 2\lfloor 2t \rfloor + 1 - i$ and $\tau(a_{2\lfloor 2t \rfloor+1-i}) = i$.
- $r = 2$: Set $\tau(a_i) = 2\lfloor 2t \rfloor + 1 - i$ and $\tau(a_{2\lfloor 2t \rfloor+1-i}) = i$.

Finally, set $\tau(a_i) = i$ for all $i > 2\lfloor 2t \rfloor$. This yields a ranking τ that is in $\mathcal{B}_t(\pi)$, and mismatches ρ in each of first $2\lfloor 2t \rfloor$ positions; hence, $d_{CY}(\rho,\tau) \geq \lfloor 2t \rfloor$, as required.

The maximum displacement distance: Finally, let d_{MD} denote the maximum displacement distance. Note that it can be at most $D = m - 1$; hence, it also has $\alpha = 1$. However, this distance metric requires an entirely different technique than the ones used for previous distances. For example, taking any two rankings at maximum distance from each other does not work. We prove this result for $t \leq m/4$. Once again, note that $d^{\downarrow}_{MD}(2t) = \lfloor 2t \rfloor$.

Consider rankings σ and σ' defined as follows.

$$\sigma = (\underbrace{a_1 \succ \ldots \succ a_{\lfloor 2t \rfloor}} \succ \underbrace{a_{\lfloor 2t \rfloor+1} \succ \ldots \succ a_{2\lfloor 2t \rfloor}} \succ a_{\text{rest}}),$$

$$\sigma' = (\underbrace{a_{\lfloor 2t \rfloor+1} \succ \ldots \succ a_{2\lfloor 2t \rfloor}} \succ \underbrace{a_1 \succ \ldots \succ a_{\lfloor 2t \rfloor}} \succ a_{\text{rest}}),$$

where a_{rest} is shorthand for $a_{2\lfloor 2t \rfloor+1} \succ \ldots \succ a_m$. Note that the blocks of alternatives a_1 through $a_{\lfloor 2t \rfloor}$ and $a_{\lfloor 2t \rfloor+1}$ through $a_{2\lfloor 2t \rfloor}$ are shifted to each other's positions in the two rankings. Thus, each of a_1 through $a_{2\lfloor 2t \rfloor}$ have a displacement of exactly $\lfloor 2t \rfloor$ between the two rankings. Thus, $d_{MD}(\sigma,\sigma') = \lfloor 2t \rfloor$.

Consider the profile π consisting of $n/2$ instances of σ and σ' each. Clearly, σ and σ' have an average distance of $\lfloor 2t \rfloor/2 \leq t$ from rankings in π. Hence, $\{\sigma,\sigma'\} \in \mathcal{B}_t(\pi)$. Surprisingly, in this case we can show that the minimax distance of $\mathcal{B}_t(\pi)$ without any additional information regarding the structure of $\mathcal{B}_t(\pi)$.

Take a ranking $\rho \in \mathcal{L}(A)$. The alternative placed first in ρ must be ranked at a position $\lfloor 2t \rfloor$ or below in at least one of σ and σ'. Hence, $\max(d_{MD}(\rho,\sigma), d_{MD}(\rho,\sigma')) \geq \lfloor 2t \rfloor$. Thus, there exists a ranking in $\mathcal{B}_t(\pi)$ at distance at least $\lfloor 2t \rfloor$ from ρ, i.e., the minimax distance of $\mathcal{B}_t(\pi)$ is at least $\lfloor 2t \rfloor$, as desired.

This completes the proof of the special case of even n for all four distance metrics. Now, consider the case of odd n.

***Odd n*:** To extend the proof to odd values of n, we simply add one more instance of σ than σ'. The key insight is that with large n, the distance from the additional vote would have little effect on the average distance of a ranking from the profile. Thus, $\mathcal{B}_t(\pi)$ would be preserved, and the proof would follow.

Formally, let $L \geq 2$ and $t \in (1/L, 1 - 1/L)$. For the case of even n, the proofs for all four distance metrics proceeded as follows: Given the feasible distance $r = d^{\downarrow}(2t)$, we constructed two rankings σ and σ' at distance r from each other such that $\mathcal{B}_t(\pi)$ is the set of rankings at minimal total distance from the two rankings, i.e.,

$$\mathcal{B}_t(\pi) = \left\{\rho \in \mathcal{L}(A) \mid d(\rho,\sigma) + d(\rho,\sigma') = r\right\}.$$

Let $n \geq 3$ be odd. Consider the profile π that has $(n-1)/2$ instances of σ and σ' each, and an additional instance of an arbitrary ranking. In our generic proof for all four distance metrics, we obtain conditions under which $\mathcal{B}_t(\pi) = \mathcal{B}_t(\pi'$

where π' is obtained by removing the arbitrary ranking from π (and hence has an even number of votes). We already proved that $k(\mathcal{B}_t(\pi')) \geq d^{\downarrow}(2t)$. Hence, obtaining $\mathcal{B}_t(\pi) = \mathcal{B}_t(\pi')$ would also show the lower bound $d^{\downarrow}(2t)$ for odd n.

In more detail, our objective is that every ranking ρ with $d(\rho,\sigma) + d(\rho,\sigma') = r$ (which may have a worst-case distance of D from the additional arbitrary ranking) should be in $\mathcal{B}_t(\pi)$, and every ranking ρ with $d(\rho,\sigma) + d(\rho,\sigma') > r$ should be outside $\mathcal{B}_t(\pi)$.

First, let $d \in \{d_{KT}, d_{CY}, d_{MD}\}$. If $d(\rho,\sigma) + d(\rho,\sigma') > r$, then $d(\rho,\sigma) + d(\rho,\sigma') \geq r + 1$. The total error incurred by rankings of distance r from π is $\frac{n-1}{2} \cdot r$, and a distance of D from the additional ranking. This means that

$$t \geq \frac{\frac{n-1}{2} \cdot r + D}{n}.$$

For rankings with an error greater than r to be outside $\mathcal{B}_t(\pi)$, we must have

$$t < \frac{\frac{n-1}{2} \cdot (r+1)}{n}.$$

Combining the inequalities, we obtain that

$$\begin{aligned} & \frac{\frac{n-1}{2} \cdot r + D}{n} \leq t < \frac{\frac{n-1}{2} \cdot (r+1)}{n} \\ \Leftrightarrow\ & \frac{n-1}{2} \cdot r + D \leq n \cdot t < \frac{n-1}{2} \cdot (r+1) \\ \Leftrightarrow\ & r + \frac{2D}{n-1} \leq \frac{2n}{n-1} \cdot t < r + 1 \\ \Leftrightarrow\ & r \leq 2t - \frac{2D-2t}{n-1} < r + \left(1 - \frac{2D}{n-1}\right). \end{aligned} \tag{9}$$

Choose $n \geq 2LD + 1$. Then, $2D/(n-1) \leq 1/L < \{2t\}$. Note that

$$\left\lfloor 2t - \frac{2D-2t}{n-1} \right\rfloor = \left\lfloor \lfloor 2t \rfloor + \{2t\} - \frac{2D-2t}{n-1} \right\rfloor = \lfloor 2t \rfloor,$$

where the last equality holds because we showed $(2D-2t)/(n-1) < \{2t\}$.

In all three distance metrics considered thus far, we had $\lfloor 2t \rfloor = d^{\downarrow}(2t)$. Let $r = \lfloor 2t \rfloor$. We show that $r = \lfloor 2t \rfloor$ satisfies Equation (9), thus yielding $\mathcal{B}_t(\pi)$ with minimax distance at least $r = d^{\downarrow}(2t)$, as required. Note that

$$r \leq 2t - \frac{2D-2t}{n-1}$$

is satisfied by definition from Equation (9). We also have

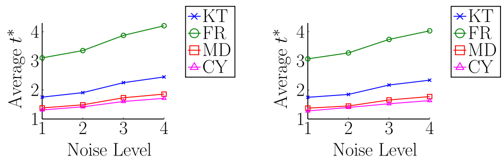
$$\begin{aligned} \left(2t - \frac{2D-2t}{n-1}\right) - \left(r + 1 - \frac{2D}{n-1}\right) &= 2t + \frac{2t}{n-1} - \lfloor 2t \rfloor - 1 \\ &= \{2t\} + \frac{2t}{n-1} - 1 \\ &< 1 - \frac{1}{L} + \frac{1}{L} - 1 = 0. \end{aligned}$$

Hence, we have $k(t) \geq d^{\downarrow}(2t)$ for $n \geq 2LD + 1$.

Next, consider the footrule distance. If $\lfloor 2t \rfloor$ is even (i.e., if $\lfloor 2t \rfloor = d^{\downarrow}(2t)$), then the above proof works because $r = \lfloor 2t \rfloor$ is a feasible distance. If $\lfloor 2t \rfloor$ is odd, then we must choose $r = \lfloor 2t \rfloor - 1$. However, we have an advantage: since the footrule distance is always even, every ranking ρ with $d(\rho,\sigma) + d(\rho,\sigma') > r$ must have $d(\rho,\sigma) + d(\rho,\sigma') \geq r + 2$. Hence, we only need

$$\begin{aligned} & \frac{\frac{n-1}{2} \cdot r + D}{n} \leq t < \frac{\frac{n-1}{2} \cdot (r+2)}{n} \\ \Leftrightarrow\ & r \leq 2t - \frac{2D-2t}{n-1} < r + \left(2 - \frac{2D}{n-1}\right). \end{aligned} \tag{10}$$

Note that $r = \lfloor 2t \rfloor - 1$ clearly satisfies the first inequality in Equation (10). For the second inequality, note that r decreased by 1 compared to earlier but $1 - 2D/(n-1)$ increased to $2 - 2D/(n-1)$ instead. Hence, the second inequality is still satisfied, and we get $\mathcal{B}_t(\pi)$ with minimax distance at least $r = \lfloor 2t \rfloor - 1 = d^{\downarrow}(2t)$, as required. $\square$



(a) Dots (b) Puzzle

Fig. 1. Positive correlation of t^* with the noise parameter.

4. Approximations for unknown average error

In the previous sections we derived the optimal rules when the upper bound t on the average error is given to us. In practice, the given bound may be inaccurate. We know that using an estimate $\widehat{t}$ that is still an upper bound ($\widehat{t} \geq t$) yields a ranking at distance at most $2\widehat{t}$ from the ground truth in the worst case. What happens if it turns out that $\widehat{t} < t$? We show that the output ranking is still at distance at most $4t$ from the ground truth in the worst case.

Theorem 6. *For a distance metric d, a profile π consisting of n noisy rankings at an average distance of at most t from the true ranking σ^*, and $\widehat{t} < t$, $d(\mathrm{OPT}^d(\widehat{t}, \pi), \sigma^*) \leq 4t$.*

To prove the theorem, we make a detour through minisum rules. For a distance metric d, let $\textsc{MiniSum}^d$, be the voting rule that always returns the ranking minimizing the sum of distances (equivalently, average distance) from the rankings in the given profile according to d. Two popular minisum rules are the Kemeny rule for the Kendall tau distance ($\textsc{MiniSum}^{d_{KT}}$) and the minisum rule for the footrule distance ($\textsc{MiniSum}^{d_{FR}}$), which approximates the Kemeny rule [18].[5] For a distance metric d (dropped from the superscripts), let $d(\pi, \sigma^*) \leq t$. We claim that the minisum ranking $\textsc{MiniSum}(\pi)$ is at distance at most $\min(2t, 2k(t, \pi))$ from σ^*. This is true because the minisum ranking and the true ranking are both in $\mathcal{B}_t(\pi)$, and Lemma 1 shows that its diameter is at most $\min(2t, 2k(t, \pi))$.

Returning to the theorem, if we provide an underestimate $\widehat{t}$ of the true worst-case average error t, then using Lemma 1,

$$d\left(\textsc{MiniMax}(\mathcal{B}_{\widehat{t}}(\pi)), \textsc{MiniSum}(\pi)\right) \leq 2\widehat{t} \leq 2t,$$
$$d\left(\textsc{MiniSum}(\pi), \sigma^*\right) \leq \mathcal{D}(\mathcal{B}_t(\pi)) \leq 2t.$$

By the triangle inequality, $d\left(\textsc{MiniMax}(\mathcal{B}_{\widehat{t}}(\pi)), \sigma^*\right) \leq 4t$.

5. Experimental results

We compare our worst-case optimal voting rules OPT^d against a plethora of voting rules used in the literature: plurality, Borda count, veto, the Kemeny rule, single transferable vote (STV), Copeland's rule, Bucklin's rule, the maximin rule, Slater's rule, Tideman's rule, and the modal ranking rule (for definitions see, e.g., [13]).

Our performance measure is the distance of the output ranking from the *actual* ground truth. In contrast, for a given d, OPT^d is designed to optimize the worst-case distance to *any possible* ground truth. Hence, crucially, OPT^d *is not guaranteed to outperform other rules in our experiments*.

We use two real-world datasets containing ranked preferences in domains where ground truth rankings exist. Mao, Procaccia, and Chen [27] collected these datasets — *dots* and *puzzle* — via Amazon Mechanical Turk. For dataset *dots* (resp., *puzzle*), human workers were asked to rank four images that contain a different number of dots (resp., different states of an 8-Puzzle) according to the number of dots (resp., the distances of the states from the goal state). Each dataset has four different noise levels (i.e., levels of task difficulty), represented using a single noise parameter: for dots (resp., puzzle), higher noise corresponds to ranking images with a smaller difference between their number of dots (resp., ranking states that are all farther away from the goal state). Each dataset has 40 profiles with approximately 20 votes each, for each of the 4 noise levels. Points in our graphs are averaged over the 40 profiles in a single noise level of a dataset.

First, as a sanity check, we verified (Fig. 1) that the noise parameter in the datasets positively correlates with our notion of noise — the average error in the profile, denoted t^* (averaged over all profiles in a noise level). Strikingly, the results from the two datasets are almost identical!

Next, we compare OPT^d and $\textsc{MiniSum}^d$ against the voting rules listed above, with distance d as the measure of error. We use the average error in a profile as the bound t given to OPT^d, i.e., we compute $\mathrm{OPT}^d(t^*, \pi)$ on profile π where $t^* = d(\pi, \sigma^*)$. While this is somewhat optimistic, note that t^* may not be the (optimal) value of t that achieves the lowest error. Also, the experiments below show that a reasonable estimate of t^* also suffices.

[5] Minisum rules such as the Kemeny rule are also compelling because they often satisfy attractive social choice axioms. However, it is unclear whether such axioms contribute to the overall goal of effectively recovering the ground truth.

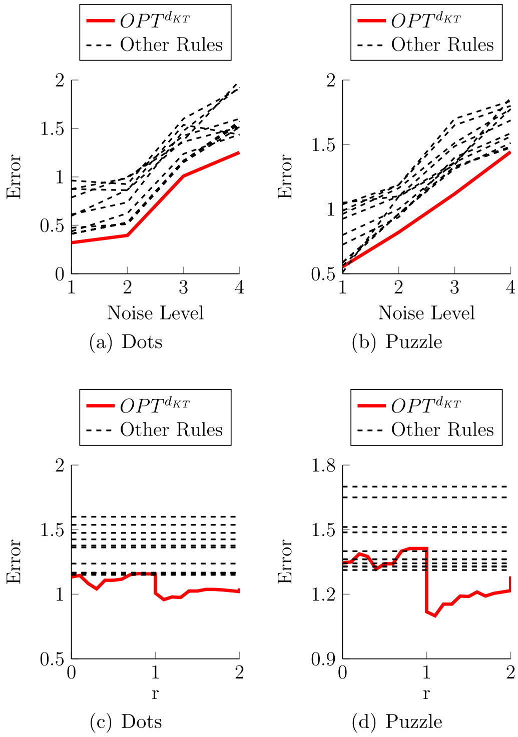


Fig. 2. Performance of different voting rules (Figs. 1(a) and 1(b)), and of OPT with varying $\widehat{t}$ (Figs. 1(c) and 1(d)). (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Figs. 2(a) and 2(b) show the results for the dots and puzzle datasets, respectively, under the Kendall tau distance. It can be seen that $\text{OPT}^{d_{KT}}$ (solid red line) significantly outperforms all other voting rules. The three other distance metrics considered in this paper generate similar results; the corresponding graphs are presented in the appendix.

Finally, we test OPT^d in the more demanding setting where only an estimate $\widehat{t}$ of t^* is provided. To synchronize the results across different profiles, we use $r = (\widehat{t} - \text{MAD})/(t^* - \text{MAD})$, where MAD is the *minimum average distance* of any ranking from the votes in a profile, that is, the average distance of the ranking returned by MiniSum^d from the input votes. For all profiles, $r = 0$ implies $\widehat{t} = \text{MAD}$ (the smallest value that admits a possible ground truth) and $r = 1$ implies $\widehat{t} = t^*$ (the true average error). In our experiments we use $r \in [0, 2]$; here, $\widehat{t}$ is an overestimate of t^* for $r \in (1, 2]$ (a valid upper bound on t^*), but an underestimate of t^* for $r \in [0, 1)$ (an invalid upper bound on t^*).

Figs. 2(c) and 2(d) show the results for the dots and puzzle datasets, respectively, for a representative noise level (level 3 in previous experiments) and the Kendall tau distance. We can see that $\text{OPT}^{d_{KT}}$ (solid red line) outperforms all other voting rules as long as $\widehat{t}$ is a reasonable overestimate of t^* ($r \in [1, 2]$), but may or may not outperform them if $\widehat{t}$ is an underestimate of t^*. Again, other distance metrics generate similar results (see the appendix for details).

Comments on the empirical results. It is genuinely surprising that on real-world datasets, OPT^d (a rule designed to work well in the worst-case) provides a significantly superior average-case performance compared to most prominent voting rules by utilizing minimal additional information — an approximate upper bound on the average error in the input votes.

The inferior performance of methods based on probabilistic models of error is also thought provoking. After all, these models assume independent errors in the input votes, which is a plausible assumption in crowdsourcing settings. But such probabilistic models typically assume a specific structure on the distribution of the noise, e.g., the exponential distribution in Mallows' model [26], and it is almost impossible that noise in practice would follow this exact structure. In contrast, our approach only requires a loose upper bound on the average error in the input votes. In crowdsourcing settings where the noise is highly unpredictable, it can be argued that the principal may not be able to judge the exact distribution of errors, but may be able to provide an approximate bound on the average error.

6. Discussion

Uniformly accurate votes. Motivated by crowdsourcing settings, we considered the case where the average error in the input votes is guaranteed to be low. Instead, suppose we know that every vote in the input profile π is at distance at most t from the ground truth σ^*, i.e., $\max_{\sigma \in \pi} d(\sigma, \sigma^*) \leq t$. If t is small, this is a stronger assumption because it means that there are no outliers, which is implausible in crowdsourcing settings but plausible if the input votes are expert opinions. In this setting, it is immediate that any vote in the given profile is at distance at most $d^{\downarrow}(t)$ from the ground truth. Moreover,

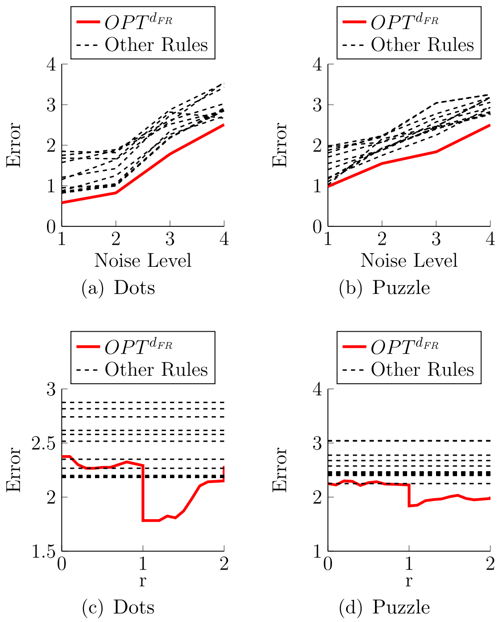


Fig. A.3. Results for the footrule distance (d_{FR}): Figs. A.3(a) and A.3(b) show that $\mathrm{OPT}^{d_{FR}}$ outperforms other rules given the true parameter, and Figs. A.3(c) and A.3(d) (for a representative noise level 3) show that it also outperforms the other rules with a reasonable estimate.

the proof of Theorem 4 goes through, so this bound is tight *in the worst case*; however, returning a ranking from the profile is *not optimal for every profile*.

Randomization. We did not consider randomized rules, which may return a distribution over rankings. If we take the error of a randomized rule to be the expected distance of the returned ranking from the ground truth, it is easy to obtain an upper bound of t. Again, the proof of Theorem 4 can be extended to yield an almost matching lower bound of $d^{\downarrow}(t)$. While randomized rules provide better guarantees, they are often impractical: low error is only guaranteed when rankings are repeatedly selected from the output distribution of the randomized rule on the same profile; however, most social choice settings see only a single outcome realized.[6]

Complexity. A potential drawback of the proposed approach is computational complexity. For example, consider the Kendall tau distance. When t is small enough, only the Kemeny ranking would be a possible ground truth, and $\mathrm{OPT}^{d_{KT}}$ or any finite approximation thereof must return the Kemeny ranking, if it is unique. The $\mathcal{NP}$-hardness of computing the Kemeny ranking [6] therefore suggests that computing or approximating $\mathrm{OPT}^{d_{KT}}$ is $\mathcal{NP}$-hard.

One way to circumvent this computational obstacle is picking a ranking from the given profile, which provides a weaker bound of $3t$ instead of $2t$ on the distance from the unknown ground truth (see Theorem 2). However, in practice the optimal ranking can also be computed using various fixed-parameter tractable algorithms, integer programming solutions, and other heuristics, which are known to provide good performance for these types of computational problems (see, e.g., [8,7]). More importantly, the crowdsourcing settings that motivate our work inherently restrict the number of alternatives to a relatively small constant: a human would find it difficult to effectively rank more than, say, 10 alternatives. With a constant number of alternatives, we can simply enumerate all possible rankings in polynomial time, making each and every computational problem considered in this paper tractable. In fact, this is what we did in our experiments. Therefore, we do not view computational complexity as an insurmountable obstacle.

Acknowledgements

This work was partially supported by the NSF under grants IIS-1350598, CCF-1215883, and CCF-1525932, and by a Sloan Research Fellowship.

Appendix A. Additional experiments

In the paper, we presented experiments (Fig. 2) that compare our proposed worst-case optimal rule against other voting rules when: i) it receives the true error of a profile $t^* = d(\pi, \sigma^*)$ as an argument (Figs. 2(a) and 2(b)), and ii) when it receives an estimate $\widehat{t}$ of t^* (Figs. 2(c) and 2(d)). In these experiments, we used the Kendall tau distance as the measure of

[6] Exceptions include cases where randomization is used for circumventing impossibilities [29,14,11].

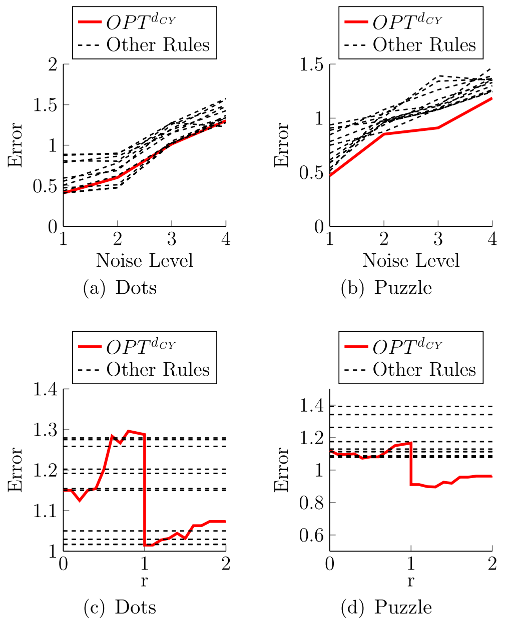


Fig. A.4. Results for the Cayley distance (d_{CY}): Figs. A.4(a) and A.4(b) show that $\text{OPT}^{d_{CY}}$ outperforms other rules given the true parameter, and Figs. A.4(c) and A.4(d) (for a representative noise level 3) show that it also outperforms the other rules with a reasonable estimate.

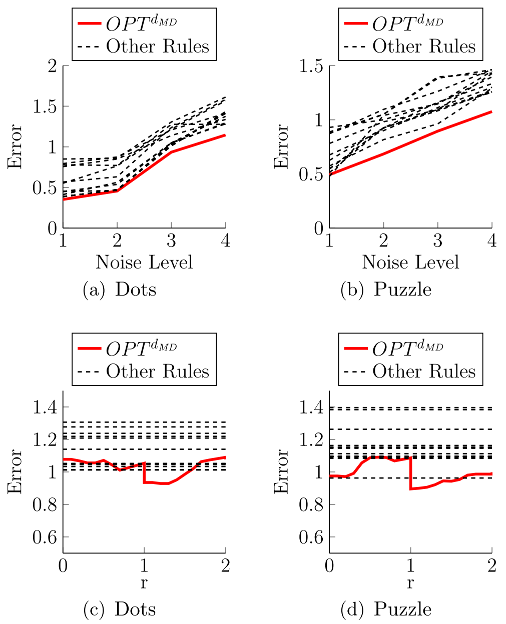


Fig. A.5. Results for the maximum displacement distance (d_{MD}): Figs. A.5(a) and A.5(b) show that $\text{OPT}^{d_{MD}}$ outperforms other rules given the true parameter, and Figs. A.5(c) and A.5(d) (for a representative noise level 3) show that it also outperforms the other rules with a reasonable estimate.

error. In this section we present additional experiments in an essentially identical setting but using the other three distance metrics considered in this paper as the measure of error. These experiments affirm that our proposed rules are superior to other voting rules independent of the error measure chosen. Figs. A.3, A.4, and A.5 show the experiments for the footrule distance, the Cayley distance, and the maximum displacement distance, respectively.

References

[1] K. Arrow, Social Choice and Individual Values, John Wiley and Sons, 1951.
[2] H. Azari Soufiani, D.C. Parkes, L. Xia, Random utility theory for social choice, in: Proceedings of the 26th Annual Conference on Neural Information Processing Systems, NIPS, 2012, pp. 126–134.

[3] H. Azari Soufiani, D.C. Parkes, L. Xia, Preference elicitation for general random utility models, in: Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence, UAI, 2013, pp. 596–605.
[4] H. Azari Soufiani, D.C. Parkes, L. Xia, Computing parametric ranking models via rank-breaking, in: Proceedings of the 31st International Conference on Machine Learning, ICML, 2014, pp. 360–368.
[5] A. Barg, A. Mazumdar, Codes in permutations and error correction for rank modulation, IEEE Trans. Inf. Theory 56 (7) (2010) 3158–3165.
[6] J. Bartholdi, C.A. Tovey, M.A. Trick, Voting schemes for which it can be difficult to tell who won the election, Soc. Choice Welf. 6 (1989) 157–165.
[7] N. Betzler, R. Bredereck, R. Niedermeier, Theoretical and empirical evaluation of data reduction for exact Kemeny rank aggregation, Auton. Agents Multi-Agent Syst. 28 (5) (2014) 721–748.
[8] N. Betzler, J. Guo, C. Komusiewicz, R. Niedermeier, Average parameterization and partial kernelization for computing medians, J. Comput. Syst. Sci. 77 (4) (2011) 774–789.
[9] A. Blum, On-Line Algorithms in Machine Learning, Springer, 1998.
[10] A. Borodin, R. El-Yaniv, Online Computation and Competitive Analysis, Cambridge University Press, 2005.
[11] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A.D. Procaccia, O. Sheffet, Optimal social choice functions: a utilitarian view, in: Proceedings of the 13th ACM Conference on Economics and Computation, EC, 2012, pp. 197–214.
[12] I. Caragiannis, A.D. Procaccia, N. Shah, When do noisy votes reveal the truth?, in: Proceedings of the 14th ACM Conference on Economics and Computation, EC, 2013, pp. 143–160.
[13] I. Caragiannis, A.D. Procaccia, N. Shah, Modal ranking: a uniquely robust voting rule, in: Proceedings of the 28th AAAI Conference on Artificial Intelligence, AAAI, 2014, pp. 616–622.
[14] V. Conitzer, Anonymity-proof voting rules, in: Proceedings of the 4th International Workshop on Internet and Network Economics, WINE, 2008, pp. 295–306.
[15] V. Conitzer, M. Rognlie, L. Xia, Preference functions that score rankings and maximum likelihood estimation, in: Proceedings of the 21st International Joint Conference on Artificial Intelligence, IJCAI, 2009, pp. 109–115.
[16] V. Conitzer, T. Sandholm, Communication complexity of common voting rules, in: Proceedings of the 6th ACM Conference on Economics and Computation, EC, 2005, pp. 78–87.
[17] Persi Diaconis, Ronald L. Graham, Spearman's footrule as a measure of disarray, J. R. Stat. Soc. B 32 (24) (1977) 262–268.
[18] C. Dwork, R. Kumar, M. Naor, D. Sivakumar, Rank aggregation methods for the web, in: Proceedings of the 10th International World Wide Web Conference, WWW, 2001, pp. 613–622.
[19] E. Elkind, P. Faliszewski, A. Slinko, Good rationalizations of voting rules, in: Proceedings of the 24th AAAI Conference on Artificial Intelligence, AAAI, 2010, pp. 774–779.
[20] V. Guruswami, List Decoding of Error-Correcting Codes, Springer, 2005.
[21] B. Haeupler, Interactive channel capacity revisited, in: Proceedings of the 55th Symposium on Foundations of Computer Science, FOCS, 2014, pp. 226–235.
[22] T.L. Heath, Diophantus of Alexandria: A Study in the History of Greek Algebra, CUP Archive, 2012.
[23] M.J. Kearns, R.E. Schapire, L.M. Sellie, Toward efficient agnostic learning, Mach. Learn. 17 (1994) 115–141.
[24] J. Lee, W. Kladwang, M. Lee, D. Cantu, M. Azizyan, H. Kim, A. Limpaecher, S. Yoon, A. Treuille, R. Das, RNA design rules from a massive open laboratory, Proc. Natl. Acad. Sci. 111 (6) (2014) 2122–2127.
[25] T. Lu, C. Boutilier, Robust approximation and incremental elicitation in voting protocols, in: Proceedings of the 22nd International Joint Conference on Artificial Intelligence, IJCAI, 2011, pp. 287–293.
[26] C.L. Mallows, Non-null ranking models, Biometrika 44 (1957) 114–130.
[27] A. Mao, A.D. Procaccia, Y. Chen, Better human computation through principled voting, in: Proceedings of the 27th AAAI Conference on Artificial Intelligence, AAAI, 2013, pp. 1142–1148.
[28] T. Meskanen, H. Nurmi, Closeness counts in social choice, in: M. Braham, F. Steffen (Eds.), Power, Freedom, and Voting, Springer-Verlag, 2008.
[29] A.D. Procaccia, Can approximation circumvent Gibbard–Satterthwaite?, in: Proceedings of the 24th AAAI Conference on Artificial Intelligence, AAAI, 2010, pp. 836–841.
[30] A.D. Procaccia, S.J. Reddi, N. Shah, A maximum likelihood approach for selecting sets of alternatives, in: Proceedings of the 28th Annual Conference on Uncertainty in Artificial Intelligence, UAI, 2012, pp. 695–704.
[31] A.D. Procaccia, J.S. Rosenschein, G.A. Kaminka, On the robustness of preference aggregation in noisy environments, in: Proceedings of the 6th International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS, 2007, pp. 416–422.
[32] L. Xia, V. Conitzer, A maximum likelihood approach towards aggregating partial orders, in: Proceedings of the 22nd International Joint Conference on Artificial Intelligence, IJCAI, 2011, pp. 446–451.
[33] L. Xia, V. Conitzer, J. Lang, Aggregating preferences in multi-issue domains by using maximum likelihood estimators, in: Proceedings of the 9th International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS, 2010, pp. 399–408.
[34] H.P. Young, Condorcet's theory of voting, Am. Polit. Sci. Rev. 82 (4) (1988) 1231–1244.